

Cheol E. Han, Luis R. Peraza, John-Paul Taylor, Marcus Kaiser

Predicting age of human subjects based on structural connectivity from diffusion tensor imaging

Technical Report No. 1

Thursday 22nd May, 2014

Dynamic Connectome Lab

<http://www.biological-networks.org/>

Predicting age of human subjects based on structural connectivity from diffusion tensor imaging

Cheol E. Han^{1,2}, Luis R. Peraza³, John-Paul Taylor³, Marcus Kaiser^{2,4,5}

¹ *Department of Biomedical Engineering, Korea University, South Korea*

² *Department of Brain Cognitive Sciences, Seoul National University, South Korea*

³ *Institute for Ageing and Health, Newcastle University, UK*

⁴ *School of Computing Science, Newcastle University, UK*

⁵ *Institute of Neuroscience, Newcastle University, UK*

E-Mail: cheolhan@gmail.com

<http://bia.korea.ac.kr/people/~cheolhan/>

Summary

Predicting brain maturity using noninvasive magnetic resonance images (MRI) can distinguish different age groups and help to assess neurodevelopmental disorders. However, group-wise differences are often less informative for assessing features of individuals. Here, we propose a simple method to predict the age of an individual subject solely based on structural connectivity data from diffusion tensor imaging (DTI). Our simple predictor computed a weighted sum of the strength of all connections of an individual. The weight consists of the fiber strength, given by the number of streamlines following tract tracing, multiplied by the importance of that connection for an observed feature—age in this case. We tested this approach using DTI data from 121 healthy subjects aged 4 to 85 years. After determining importance in a training dataset, our predicted ages in the test dataset showed a strong correlation ($\rho = 0.77$) with real age deviating by, on average, only 10 years.

1 Introduction

The study of how different components of the brain, may they be neurons or brain regions, are connected has become an emerging field within the neurosciences [1, 2, 3]. Structural connectivity observes the physical wiring of neural circuits, while functional connectivity links brain regions with similar activity over time. Magnetic Resonance Imaging (MRI) facilitates human studies since it enables us to construct such networks noninvasively. At the macro scale where network nodes represent brain regions, Diffusion Tensor Imaging (DTI) allows us to quantify the number of streamlines after tract tracing as a proxy of connection strength between two regions. Analyzing brain networks can be a tool to understand the interaction between nodes and has already shown a strong correlation between cognitive functioning and global information integration for functional networks [4]. Predicting brain maturity from MR images is beneficial to assess neurodevelopmental disorders. As pediatric disorders including ADHD have delays in brain maturity [5], investigating structural normality of pediatric brains can be a good screening strategy. Also in the elderly, it is beneficial to distinguish cognitive disorders from normal ageing. Taking advantage of the network analysis framework, classifying structural connectivity networks to distinguish different age ranges has previously been performed by looking at small sets of nodes, network motifs [6], or by sets of features for individual nodes, single-node motifs [7, 8]. Also network topological changes over age were investigated in structural connectivity [9, 10]. Whereas these studies use aggregate features of the network, we used only raw data about individual edges, as given by the number of streamlines following tract tracing of DTI data, to inform predictions of the age of a subject. While a link between functional connectivity and brain maturity was reported earlier [11], it remains unclear whether structural connections alone can be a comparatively suitable predictor as functional and structural connectivity are often related [12]. Here, we tested a simple prediction model for the age of a subject using information about structural connectivity based on diffusion tensor imaging. Our predictor computed a weighted sum of the structural connectivity matrix for each subject as a raw score, where the weight of each connection was predefined as the correlation coefficient between connection strength of that connection with age over all subjects in the training group. Then, we transformed the raw score to a predicted age with linear regression. We then evaluated our method over a test group, computing the correlation between predicted ages and the real ages in the test group.

2 Methods

2.1 DTI database

We made use of a public DTI-database (<http://fcon.1000-projects.nitrc.org/indi/pro/nki.html>) provided by the Nathan Kline Institute (NKI) [13]. DTI-data were obtained with a 3 Tesla scanner (Siemens MAGNETOM TrioTim syngo, Erlangen, Germany). T1 weighted MRI data were obtained with 1mm isovoxel, FoV 256mm, TR=2500ms, and TE=3.5ms. DTI data were recorded with 2mm isovoxel, FoV=256mm, TR=10000ms, TE=91ms, and 64 diffusion directions with b-factor of 1000 s mm⁻² and 12 b₀ images. We included 121 participants between 4 and 85 years. We used Freesurfer to obtain surface meshes of the boundary between gray and white matter from T1 anatomical brain images (<http://surfer.nmr-mgh.harvard.edu>), see Figure 1. After registering surface meshes into the DTI space, we generated volume regions of interest (ROIs) based on GM-voxels. Freesurfer provides parcellation of 34 cortical regions based on the Desikan atlas [16] and 7 subcortical regions (Nucleus accumbens, Amygdala, Caudate, Hippocampus, Pallidum, Putamen, and Thalamus) [17] for each hemisphere, thus leading to 82 ROIs in total. To quantify the connection strength from DTI, we performed eddy-current correction (FSL), and the Fiber Assignment by Continuous Tracking (FACT) algorithm [14] with 35 degrees angular threshold using Diffusion Toolkit along with TrackVis [15]. We then counted the number of streamlines between all pairs of defined regions of interest as connection strength

using the UCLA Multimodal Connectivity Package (<http://ccn.ucla.edu/wiki/index.php>). This led to an undirected weighted connectivity matrix \mathbf{S} of size $N \times N$ with $N(N-1)/2$ connections per subject, where N is 82. We also defined \mathbf{L} as a $N \times N \times M$ matrix with the connectivity matrices \mathbf{S}_k of $M = 121$ subjects, where k is the subject index.

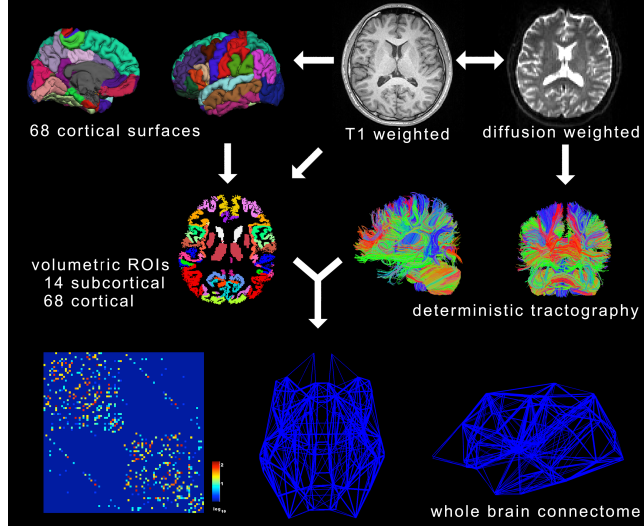


Figure 1: DTI network inference.

2.2 Age prediction

Numerous machine-learning approaches could be applied for predicting features such as age. Instead, we focused on simple predictions that look at the correlation between the strength of each fiber tract, as measured by the number of streamlines and age. The correlation, ranging from -1 to 1, gives us a weight measure of how important each connection is for a feature leading to the group correlation matrix \mathbf{C} . The stronger a connection, given by the number of streamlines, the larger the impact on the predicted feature. Correlation matrix \mathbf{C} is defined as $C_{ij} = cov(L_{ij}, A) / (\sigma_L \sigma_A)$, where L_{ij} is the vector of connection strength between node i and j for all subjects, A is the vector of ages ordered according to the M subjects, and cov and σ stands for covariance and variance respectively.

In order to predict age, a predictor value P is then given by the sum over all edges of the individual product between the correlation matrix \mathbf{C} and the subject's matrix \mathbf{S} as follows;

$$P_k = \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N S_{ij} \times C_{ij} . \quad (1)$$

To compute predictor P , we divided our dataset in two groups; a training matrix \mathbf{L}_R and a test matrix \mathbf{L}_T , where the former was used to compute \mathbf{C} and the latter to compute P . Using the training data set, P was then translated and re-scaled to map the age variable. Additionally, we tested different ways of normalizing the group matrices \mathbf{L}_R and \mathbf{L}_T ; 1) No normalization, \mathbf{L} , 2) normalization between subjects, \mathbf{L}' , 3) normalization within subjects, \mathbf{L}'' . The last two are defined as

$$L'_{ijk} = \frac{L_{ijk}}{\frac{1}{M} \sum_{i=j}^M L_{ijk}} , \quad (2)$$

$$L''_{ijk} = \frac{L_{ijk}}{\frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N L_{ijk}} . \quad (3)$$

Notice that in the previous equations M is the number of subjects in each of the groups, and normalization is applied separately to \mathbf{L}_R and \mathbf{L}_T .

2.3 Simulations

To test our age prediction, we used a bootstrapping approach. Testing \mathbf{L}_T and training \mathbf{L}_R matrices were created by randomly dividing the population matrix \mathbf{L} at 50% (half of the population for training or test), and then correlation matrix \mathbf{C} and prediction vector P for each division were computed. We used 100 divisions to obtain extreme and average performances of our age prediction.

3 Results

Figure 2 showed the results for age prediction experiments. For these, the entire database was divided in two matrices: training (60 subjects) and testing (61 subjects) matrices whose subjects were chosen randomly 100 times. For each of the 100 prediction iterations an age-edge correlation matrix \mathbf{C} and predictor vector P were computed. Figure 2A showed the correlation coefficient distribution between the predicted ages and the real ages for the 100 experiments; the median was $\rho = 0.774$ (see also Table 1) with $\rho = 0.68$ and $\rho = 0.81$ as worst and best-case predictions, respectively. Figure 2B showed the best case prediction from the 100 experiments and the translation/re-scaling mapping equation of predictor vector P for age.

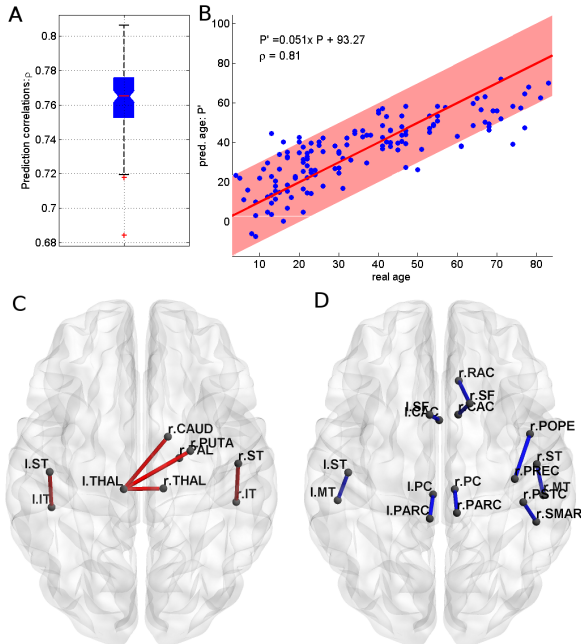


Figure 2: Age prediction results. A) Correlation box plot of 100 prediction experiments by randomly dividing the population matrix in training and test datasets. B) Best case from the 100 experiments where correlation between predicted and real ages is $\rho = 0.81$, with 95% confidence interval shown as shades. C) Highest 1% of the edges with positive correlation with age. D) Lowest 1% of the edges with negative correlation with age. Correlated edges were visualized with BrainNet Viewer (<http://www.nitrc.org/projects/bnv/>).

Prediction results for the three normalization methods tested were shown in Table 1. The

Table 1: Performance of age prediction. The best performance was obtained by non-normalization of the connectivity matrix. The table shows the first, second, and third quartiles from the correlation distribution of 100 prediction experiments for each normalization method.

| Normalization | Q_1 | Q_2 | Q_3 |
|---------------|--------|--------------|--------|
| L | 0.7603 | 0.774 | 0.7878 |
| L' | 0.5604 | 0.599 | 0.6289 |
| L'' | 0.6496 | 0.678 | 0.7095 |

best performance was obtained without normalization, followed by normalization within subjects L'' .

To infer which brain edges were most correlated with age prediction, the 100 estimated correlation matrices \mathbf{C} were averaged and the most important edges with positive and negative correlations were extracted. Figure 2C showed the 1% edges with the highest correlation with age. Similarly, Figure 2D showed the 1% of the network edges with lowest negative correlation with age. These edges and their averaged correlations were listed in Table 2.

4 Discussion

Our simple predictor estimated age of a subject given structural connectivity data with fair performance. The predicted age was highly correlated with the real age (best $\rho = 0.81$). The top edges correlated with age included the previously reported areas: Regional efficiency, capturing regional information integration, was correlated with age positively in temporal and frontal regions and negatively in parietal and occipital regions previously [9]. Though it was hard to compare directly due to different atlases and different modalities, they also found superior and inferior temporal gyri, and superior frontal gyri as we did. The edges with negative correlations were partially matched with white matter integrity (Fractional Anisotropy, FA) changes over age in a voxel based DTI study [18]: notably, both superior frontal gyri, superior temporal gyri, anterior and posterior cingulate cortices. We note that to match ours with their results, more extensive investigation is required, because we only showed the top 1% of edges.

Our method is quite simple but connected with machine learning concepts. The concept behind our method is the basic artificial neural networks, especially a radial basis function (RBF) network. Each RBF responded to a corresponding edge, and its learnable weight is a correlation coefficient we computed, capturing its influence on age. Statistically, this idea can be connected with the partial least square (PLS) regression [19]. PLS projects the original predictors (edges) into a lower dimensional space, using colinearity of each predictor variable with the feature variable (age) to be fitted. In our method, a simple Pearson correlation coefficient, that represents fitting quality of linear regression, captured colinearity. Then, a weighted sum of a subject’s connection strength represented projection of the subject into the predicting score P , where the vector of correlation coefficients was projection mapping, while PLS combines them in a statistically meaningful way.

We observed that non-normalization performed best in general, while between-subject normalization deteriorated its performance. First of all, we noted that this normalization was not standardization with the mean and standard deviation of each edge, but regularization of magnitudes to minimize effects of variability in edges’ magnitudes. The vector of correlation coefficients captured the pattern of edges correlated with age. Because correlation coefficients captured the fitting quality (a clear relationship), not the magnitude of influence itself, we worried that a few edges whose overall magnitude is relatively larger than for other edges dominantly affected the prediction. So, we regularized edges with large average magnitudes by dividing the group mean of the edge (between-group normalization), and the individual sum of edges (within-group normalization). Though the between-group normalization does not affect correlation coefficients,

Table 2: List of strongest top 1% connections correlated with age

| Edge | $\bar{\rho}$ |
|--|--------------|
| Positive correlations | |
| L. Thalamus - R. Thalamus. | 0.3562 |
| L. Superior temporal - L. Inferior temporal gyrus. | 0.3534 |
| L. Thalamus - R. Caudate Nucleus. | 0.3135 |
| L. Thalamus - R. Putamen. | 0.3018 |
| L. Thalamus - R. Pallidum. | 0.2872 |
| R. Superior temporal - R. Inferior temporal gyrus. | 0.2902 |
| Negative correlations | |
| L. Posterior cingulate cortex - L. Paracentral gyrus. | -0.5905 |
| L. Superior frontal - L. Caudal anterior cingulate cortex. | -0.5613 |
| R. Superior temporal - R. Middle temporal gyrus. | -0.5340 |
| L. Superior temporal - L. Middle temporal gyrus. | -0.5204 |
| R. Posterior cingulate cortex - R. Paracentral gyrus. | -0.4797 |
| R. Superior frontal - R. Caudal anterior cingulate cortex. | -0.4791 |
| R. Parso percularis - R. Precentral gyrus. | -0.4518 |
| R. Rostral anterior cingulate - R. Superior frontal gyrus. | -0.4373 |
| R. Postcentral gyrus - R. Supramarginal gyrus. | -0.4324 |

it performed worse than the within-group normalization, which deteriorated correlation. Thus, we believed that not only the clear relationship captured by the correlation coefficients, but also the magnitude of the edge was important. In other words, edges with larger magnitudes (preponderantly larger variance) had greater roles in predicting brain maturity than edges with smaller magnitudes.

Reducing the number of features is crucial in machine learning literature, including recursive feature elimination (RFE) [20]. Though we did not remove any edges, when its correlation coefficient is near zero, the edge can be considered as a removed feature. Using correlation coefficients for feature elimination was common [11, 22]. Dosenbach et al. [11] selected connections with 200 largest absolute values of correlation coefficients to predict age, which survived after Bonferroni correction. We note that such a selection with Bonferroni did not work for our dataset; we may employ a (less conservative) cluster-based correction method for correlation coefficients [21]. A notably similar work with DTI and lifespan age prediction [20] used various white matter parameters including white matter integrity of voxels and support vector regression with RFE. Though the model showed slightly stronger correlation between the predicted age and the real age ($\rho \approx 0.89$), our method has the advantage of providing direct linkages between regions, which are useful to understand dynamic changes of brain connectivity during development.

The method can be improved in various ways. First, the use of nonlinear fitting may help. A potential limitation is the use of a linear fit for converting the raw prediction measure P to the age prediction, P' . While a linear fit yielded good results, exponential, logarithmic or other fits might be more suitable. For example, we noticed an underestimation of age for subjects older than 50 years which links with earlier studies that found nonlinear trends during development for brain volume [23, 24], white matter integrity [23, 24, 20], and functional connectivity [11]. Second, we may use partial correlation coefficients [22] to adjust group biases, or Spearman correlation coefficient to handle non-normality of connection strengths. Third, separating males and females may improve the performance as they may follow different developmental trajectories [9, 25].

5 Conclusion

Using a simple predictor based on the influence of each fiber tract connection on a feature of interest, we were able to predict age of a subject given structural connectivity data with fair performance. This indicates that predicting other cognitive and general features (e.g. gender) might be derived from diffusion tensor imaging data of a subject. Also the proposed method is applicable to networks with other modalities including functional connectivity. We note that it can be improved in various ways still, including nonlinear mapping between a raw prediction score P and a predicted age P' , and the use of partial correlation coefficients.

Besides the clinical and basic scientific applications, our results also raise an ethical issue. If anonymized DTI data of a subject and the metadata table of name, IQ, and age of all subjects is available, one might identify the name of the subject by finding table entries that match the predicted IQ and age. In other words, even without a key (identifier) that links entries of the MRI database with those of the clinical scores/metadata database, anonymized data could be linked to a specific person.

References

- [1] O. Sporns, et al. “Organization, development and function of complex brain networks” *Trends Cogn Sci*, vol. 8, pp. 418-25, 2004.
- [2] E. Bullmore and O. Sporns “Complex brain networks: graph theoretical analysis of structural and functional systems” *Nat Rev Neurosci*, vol. 10, pp. 186-98, Mar 2009.
- [3] M. Kaiser “A Tutorial in Connectome Analysis: Topological and Spatial Features of Brain Networks” *Neuroimage*, vol. 57, pp. 892-907, 2011.
- [4] M.P. van den Heuvel, et al. “Efficiency of Functional Brain Networks and Intellectual Performance” *Journal of Neuroscience*, vol. 29, pp. 7619-7624, 2009.
- [5] P. Shaw, et al. “Development of Cortical Surface Area and Gyrification in Attention-Deficit/Hyperactivity Disorder” *Biological Psychiatry*, vol. 72, No. 3, pp. 191-7, 2012.
- [6] P. Ribeiro, et al. “Strategies for network motifs discovery” *Fifth IEEE International Conference on e-Science*, pp. 80-87, 2009.
- [7] L. Costa, et al. “Beyond the average: Detecting global singular nodes from local features in complex networks” *EPL (Europhysics Letters)*, vol. 87, p. 18008, 2009.
- [8] C. Echtermeyer, et al. “Integrating temporal and spatial scales: Human structural network motifs across age and region-of-interest size” *Front Neuroinformatics*, vol. 5, no. 10, Jul 2011.
- [9] G. Gong et al. “Age- and Gender-Related Differences in the Cortical Anatomical Network” *The Journal of Neuroscience*, vol. 29, no. 50, pp. 15684-15693, December 2009
- [10] P. Hagmann et al. “White matter maturation reshapes structural connectivity in the late developing human brain” *PNAS*, vol. 107, no. 44, pp. 19067-72, Oct 2010
- [11] N.U.F. Dosenbach, et al. “Prediction of Individual Brain Maturity Using fMRI” *Science*, vol. 329, pp. 1358-1361, 2010.
- [12] C.J. Honey, et al. “Predicting human resting-state functional connectivity from structural connectivity” *PNAS*, vol. 106, no. 6, pp. 2035-2040, 2009.
- [13] K.B. Noonan, et al. “The NKI-Rockland Sample: A Model for Accelerating the Pace of Discovery Science in Psychiatry” *Frontiers in Neuroscience*, vol. 6, no. 152, 2012.

- [14] S. Mori, and P.B. Barker “TDiffusion magnetic resonance imaging: its principle and applications” *Anat Rec*, vol. 257, pp. 102-109, 1999.
- [15] R. Wang et al. “Diffusion Toolkit: A Software Package for Diffusion Imaging Data Processing and Tractography” *In. Proc. Intl. Soc. Mag. Reson. Med.*, p 3720, 2008.
- [16] R.S. Desikan, et al. “An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest” *Neuroimage*, vol. 31, pp. 968-80, Jul 2006.
- [17] B. Fischl, et al. “Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain” *Neuron*, vol. 33, pp. 341-55, Jan 2002.
- [18] A. Stadlbauer et al. “Magnetic resonance fiber density mapping of age-related white matter changes” *European Journal of Radiology*, vol. 81, pp. 4005-4012, 2012
- [19] S. Wold et al. “The collinearity problem in linear regression. the partial least squares (PLS) approach to generalized inverses” *SIAM Journal on Scientific and Statistical Computing* , vol. 5, no. 3, pp.735-743. 1984
- [20] B. Mwangi et al. “Prediction of individual subject’s age across the human lifespan using diffusion tensor imaging: A machine learning approach” *Neuroimage*, vol. 75, pp. 58-67, 2013
- [21] C.E. Han et al. “Cluster-Based Statistics for Brain Connectivity in Correlation with Behavioral Measures” *PLoS One*, vol. 8, no. 8, e72332. 2013
- [22] R. Yun et al. “A Prediction Model for Cognitive Performance in Health Ageing Using Diffusion Tensor Imaging with Graph Theory” *35th Annual International Conference of the IEEE EMBS Osaka, Japan*, Jul 2013
- [23] L.T. Westlye, et al. “Life-span changes of the human brain white matter: diffusion tensor imaging (DTI) and volumetry” *Cereb Cortex*, vol. 20, no. 9, pp. 2055-68. Sep 2010
- [24] C. Lebel, et al. “Diffusion tensor imaging of white matter tract evolution over the lifespan” *Neuroimage*, vol. 60, no. 1, pp. 340-52. Mar 2012
- [25] S. Lim, et al. “Preferential detachment during human brain development: Age- and sex-specific structural connectivity in Diffusion Tensor Imaging (DTI)” *Cerebral Cortex*, Advance Access Dec 2013